Specification #1: Difference Map Calculation

Version: 0.1 15/SEP/09

Changes

0.1: Initial revision, will define the current state of play based on feedback from Dave Brown of Pfizer.

Purpose

To define the steps involved in the calculation of difference maps for e.g. ligand binding studies. This will involve interaction with the data reduction, beamline information management and the user.

Preconditions

The following preconditions are necessary:

- The correct unit cell constants and spacegroup are known
- A previously refined model is available from which to calculate phases

All things being equal it may be assumed that the correct cell constants correspond to those from data reduction, and that the spacegroup from the previously refined model is correct. If the cell constants make it obvious what the correct setting is (for P21212 for example) then it may be necessary to reindex the measurements. In some cases there may be ambiguity in the definition of origin, e.g. P4. In such cases it will be necessary to determine the correct origin definition for the data or model.

Postconditions

A successful outcome will be to have an electron density map which may be directly compared with the original refined model to look for additional unmodelled density. In general the experimenter will know what they are looking for and where to look for it, so getting to this point *with no input* would represent a substantial saving in effort.

N.B. that the success of the process is that a useful map has been calculated from the available data: if there is no ligand visible or present then this is a successful result provided that it is correct.

Output should consist of phased reflection file, refined model and refinement residuals. Full detail of this to be defined.

Error States

The following error states may be defined:

- Data processed with wrong unit cell.
- Model cell / symmetry not compatible with measured data.

Data Processed with Incorrect Cell

I.e. the data reduction process obtained the incorrect cell. This may be avoided by providing the correct cell and symmetry to the data reduction software. It may however be the case that the crystal symmetry is indeed different. In this case the second error state will be appropriate.

Model not Compatible with Data

Raise exception. This may have resulted from data management errors further up the stack so it is important that the user is aware of this: Will need to be able to record this error state in the data base.

Process

In the first pass the following procedure will be implemented:

- Prepare intensity data
- Determine correct setting for the data
- Perform rigid body refinement to allow for small changes in the molecule orientation
- Perform restrained refinement to calculate map from resulting reorientated molecule.

The data preparation step will be defined in a separate use case, as this will be used in several places (this will essentially take the intensity measurements from data reduction and as necessary assign the correct spacegroup, then calculate structure factor amplitudes using the TRUNCATE procedure and add a FreeR column. It may be necessary to allow provision of a previous data set which may already have FreeR assigned, in which case complete this. This should include an analysis of the axial reflections to estimate the correct spacegroup screw axes.

To determine the correct setting from a model alone it will be necessary to compare the cell constants if these are not ambiguous i.e. differences between a, b, c more than 10 percent. If there is some ambiguity then it will be necessary to calculate a reference reflection file using sfcalc, sftools and cad, then use pointless to calculate the appropriate reindexing operation to this reference for the measured diffraction data.

The refinement steps will use refmac5 and a fairly standard set of scripts. This is necessary as many industrial users are not licensed to use phenix and I also don't know where we stand with this.

Licensing

During development this will not be distributed. When it is "finished" the resulting pipeline should be made available to CCP4 using an EDNA framework to provide the interface.